Risk and Portfolio Management
Spring 2010

Auto-regressive Models
ARCH(p), GARCH(p,q)
Following R. Engle and T. Bollerslev
Conditional Mean and Conditional Variance

\(y_t, \quad t = 1, 2, 3, ..., T\)  

Given time series

\[p(y_t \mid y_{t-1}, y_{t-2}, ...) = p(y_t \mid \Phi_{t-1})\]  

Model the conditional distributions

\[y_t = \mu(\Phi_{t-1}) + \sigma(\Phi_{t-1})\varepsilon_t, \quad E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = 1\]

Example:

\[y_t \mid \Phi_{t-1} \sim N(\mu(\Phi_{t-1}), \sigma^2(\Phi_{t-1}))\]
**ARCH(p) (Engle, 1982)**

Unlike in AR, the error is not assumed to have constant variance.

More generally,

\[
    h_t = a_0 + \sum_{k=1}^{p} a_k u_{t-k}^2
\]

Uncorrelated residuals does not necessarily imply independent residuals.

Conditional variance is a lagged sum of squared residuals, eg.

\[
    h_t = \frac{1}{T} \sum_{k=1}^{T} u_{t-k}^2
\]
GARCH(p,q) (Bollerslev, 1986)

\[ u_t = h_t^{1/2} \varepsilon_t \]
\[ E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = 1 \]

\[ h_t = \omega + \sum_{i=1}^{p} \alpha_i u_i^2 + \sum_{j=1}^{q} \beta_j h_{t-j} \]

Dependence on previous squared returns and previous conditional variances.

Most famous versions in practice: GARCH(1,1) or GARCH (1,p) which are basically AR(p) processes on the conditional variance driven by the squared-returns process.
GARCH(1,1) is an exponentially weighted moving average of squared-errors. Beta determines the effective "window size" for estimation of conditional variance.
GARCH(1,2)

\[
\begin{pmatrix}
    h_t \\
    h_{t-1}
\end{pmatrix} = 
\begin{pmatrix}
    \omega \\
    0
\end{pmatrix} + \begin{pmatrix}
    \alpha & 0 \\
    0 & 0
\end{pmatrix} \begin{pmatrix}
    u_{t-1}^2 \\
    0
\end{pmatrix} + \begin{pmatrix}
    \beta_1 & \beta_2 \\
    1 & 0
\end{pmatrix} \begin{pmatrix}
    h_{t-1} \\
    h_{t-2}
\end{pmatrix}
\]

Vector AR(1)

Stability condition: \[ \lambda^2 - \beta_1 \lambda - \beta_2 = 0 \Rightarrow |\lambda| < 1 \]

\[
h_t = h + A \sum_{k=1}^{\infty} \lambda_1^k u_{t-k}^2 + B \sum_{k=1}^{\infty} \lambda_2^k u_{t-k}^2
\]

Steady-state solution

Intuitively, GARCH(1,2) is the sum of two EWMA with different time-scales (decay rates).

Notice however that the right-hand side depends on \( h \) as well, so the PDF of the conditional variance is not a chi-squared.

GARCH(1,p) is the sum of (at most) p EWMAs.
Returns of S&P 500 Index
12/1/2000-2/26/2010
Fitting to GARCH(1,p)

We know that the tails of SPY are heavy and behave like Student t with df~3.5

This heavy-tailed behavior of stock prices can be modeled by assuming a static distribution (Student) or a time-dependent distribution with a GARCH-type stochastic conditional variance.

The latter approach (GARCH) has the advantage that it incorporates dynamics so it may capture "persistence" of volatility across time.

From a portfolio risk-management perspective, the situation is "cured" by assuming a Student-t distribution with 3.5 degrees of freedom for returns (to capture tail behavior) and an EWMA variance which is adjusted daily to capture volatility clustering effects.

The question that remains is: what is the correct estimation window?
GARCH(1,1) estimation of SPY returns

Method: ML - BFGS with analytical gradient
date: 03-02-10
time: 18:10
Included observations: 2320
Convergence achieved after 56 iterations

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>omega</td>
<td>2.85989E-06</td>
<td>3.9342E-07</td>
<td>7.269290633</td>
</tr>
<tr>
<td>alpha_1</td>
<td>0.698241421</td>
<td>0.020073908</td>
<td>34.78353205</td>
</tr>
<tr>
<td>beta_1</td>
<td>0.508888808</td>
<td>0.050794297</td>
<td>10.01862092</td>
</tr>
</tbody>
</table>

Log Likelihood 7053.473574
Jarque Bera 12844.90612
Ljung-Box 65535
GARCH(2,1) estimation

Method: ML - BFGS with analytical gradient

date: 03-03-10
time: 13:25
Included observations: 2320
Convergence achieved after 45 iterations

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>omega</td>
<td>2.69557E-05</td>
<td>2.4E-06</td>
<td>11.25236</td>
<td>0</td>
</tr>
<tr>
<td>alpha_1</td>
<td>0.541398855</td>
<td>0.073788</td>
<td>7.337198</td>
<td>2.1805E-13</td>
</tr>
<tr>
<td>alpha_2</td>
<td>0.355438292</td>
<td>0.035892</td>
<td>9.90302</td>
<td>0</td>
</tr>
<tr>
<td>beta_1</td>
<td>0.268210539</td>
<td>0.045356</td>
<td>5.913404</td>
<td>3.3511E-09</td>
</tr>
</tbody>
</table>

Log Likelihood: 7060.668319

Jarque Bera: 12844.90612  Prob: 0

Ljung-Box: 65535  Prob: 65535
The document contains the results of a GARCH(1,2) model estimation. The method used was ML-BFGS with analytical gradient. The estimation was performed on March 3, 2010, at 13:34. The model was estimated using 2320 observations, and convergence was achieved after 54 iterations.

The estimated coefficients, standard errors, z-statistics, and p-values are as follows:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omega</td>
<td>1.93253E-06</td>
<td>3.45079E-07</td>
<td>5.600257981</td>
<td>2.14033E-08</td>
</tr>
<tr>
<td>Alpha_1</td>
<td>0.347594236</td>
<td>0.053959618</td>
<td>6.441747563</td>
<td>1.18106E-10</td>
</tr>
<tr>
<td>Beta_1</td>
<td>0.417978993</td>
<td>0.040988575</td>
<td>10.19745117</td>
<td>0</td>
</tr>
<tr>
<td>Beta_2</td>
<td>0.329591408</td>
<td>0.064169394</td>
<td>5.136271201</td>
<td>2.80243E-07</td>
</tr>
</tbody>
</table>

The log likelihood is 7119.174476. The Jarque Bera test statistic is 12844.90612, with a p-value of 0. The Ljung-Box test statistic is 65535, with a p-value of 65535.
Which model should we use?

All three GARCH models fit the data very well, with high z-statistics.

Preference should be given to the model with smallest number of parameters, so GARCH(1,1) should be suitable.
Cointegration and Pairs Trading

\[ X_t = \text{return on XLK} \]
\[ Y_t = \text{return on EBAY} \]

Perform \( m \)-day regression to construct residuals

\[ Y_t = \beta X_t + \varepsilon_t \]

\[ \beta = \text{SLOPE}\left((Y_{t-m},...,Y_{t-1}),(X_{t-m},...,X_{t-1})\right) \]

\[ \varepsilon_t = Y_t - \beta X_t \]

\[ \text{P & L} = 100 \prod_{k=1}^{t} (1 + \varepsilon_k) \quad y_t = y_0 + \sum_{k=1}^{t} \ln(1 + \varepsilon_k) \]

Question of interest: is \( y_t \) stationary? Does \( y_t \) have a `unit root`?
Dickey-Fuller Test for Unit Roots (aka Augmented Dickey-Fuller test)

The Dickey-Fuller test is used to test for unit roots in statistical data.

Consider the following model for the differentiated time-series:

$$\Delta y_t = \alpha + \beta t + \delta_0 y_{t-1} + \sum_{k=1}^{n} \delta_k \Delta y_{t-k} + \varepsilon_t, \quad \Delta y_t = y_t - y_{t-1}$$

Null hypothesis: there is a unit root, i.e. \( \delta_0 = 0 \).

\[
DF = \frac{\hat{\delta}_0}{\text{stdev}(\hat{\delta}_0)}
\]

ADF Critical Values:
- Reject delta=0 if DF <
  - 1% level: -3.970385
  - 5% level: -3.415895
  - 10% level: -3.130187

\( n \) is determined dynamically as part of the test (Akaike Information Criterion)
EBAY vs. XLK residuals

\[ Y_t = \text{daily return of EBAY} \]

\[ X_t = \text{daily return of XLK} \]

\[ \epsilon_t = Y_t - \beta(t-1,t-60) \cdot X_t \]

\[ y_t = y_0 + \sum_{k=1}^{t} \ln(1 + \epsilon_k), \quad y_0 = 100 \]
Augmented DF test for EBAY/XLK

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>tseries(-1)</td>
<td>-0.025582</td>
<td>0.009132</td>
<td><strong>-2.801401</strong></td>
<td>0.005222</td>
</tr>
<tr>
<td>D(tseries(-1))</td>
<td>-0.104975</td>
<td>0.036984</td>
<td><strong>-2.838362</strong></td>
<td>0.004660</td>
</tr>
<tr>
<td>D(tseries(-2))</td>
<td>0.032844</td>
<td>0.037145</td>
<td><strong>0.884206</strong></td>
<td>0.376875</td>
</tr>
<tr>
<td>D(tseries(-3))</td>
<td>0.041696</td>
<td>0.036765</td>
<td><strong>1.134124</strong></td>
<td>0.257113</td>
</tr>
<tr>
<td>D(tseries(-4))</td>
<td>-0.139433</td>
<td>0.036498</td>
<td><strong>-3.820306</strong></td>
<td>0.000145</td>
</tr>
<tr>
<td>D(tseries(-5))</td>
<td>0.023322</td>
<td>0.036852</td>
<td><strong>0.632844</strong></td>
<td>0.527033</td>
</tr>
<tr>
<td>D(tseries(-6))</td>
<td>-0.103297</td>
<td>0.036384</td>
<td><strong>-2.839106</strong></td>
<td>0.004649</td>
</tr>
<tr>
<td>D(tseries(-7))</td>
<td>-0.123580</td>
<td>0.036566</td>
<td><strong>-3.379630</strong></td>
<td>0.000764</td>
</tr>
<tr>
<td>D(tseries(-8))</td>
<td>0.062589</td>
<td>0.036842</td>
<td><strong>1.698850</strong></td>
<td>0.089771</td>
</tr>
<tr>
<td>D(tseries(-9))</td>
<td>0.103669</td>
<td>0.036604</td>
<td><strong>2.832135</strong></td>
<td>0.004751</td>
</tr>
<tr>
<td>C</td>
<td>0.120657</td>
<td>0.043010</td>
<td><strong>2.805296</strong></td>
<td>0.005160</td>
</tr>
<tr>
<td>@trend</td>
<td>-0.000006</td>
<td>0.000003</td>
<td><strong>-2.076142</strong></td>
<td>0.038228</td>
</tr>
</tbody>
</table>

Best lag fit: 9

Cannot reject UR @ 90% level
EBAY vs. QQQQ residuals
Null Hypothesis: tseries has a unit root

Exogenous: Constant and linear Trend

Lag Length: 4 (Automatic Based on AIC, MAXLAG=10)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>tseries(-1)</td>
<td>-0.023280</td>
<td>0.008338</td>
<td>-2.791940</td>
<td>0.005374</td>
</tr>
<tr>
<td>D(tseries(-1))</td>
<td>-0.078624</td>
<td>0.036419</td>
<td>-2.158873</td>
<td>0.031179</td>
</tr>
<tr>
<td>D(tseries(-2))</td>
<td>0.019488</td>
<td>0.036533</td>
<td>0.533428</td>
<td>0.593897</td>
</tr>
<tr>
<td>D(tseries(-3))</td>
<td>0.030306</td>
<td>0.036525</td>
<td>0.829726</td>
<td>0.406960</td>
</tr>
<tr>
<td>D(tseries(-4))</td>
<td>-0.114959</td>
<td>0.036359</td>
<td>-3.161785</td>
<td>0.001632</td>
</tr>
<tr>
<td>C</td>
<td>0.109870</td>
<td>0.039251</td>
<td>2.799187</td>
<td>0.005256</td>
</tr>
<tr>
<td>@trend</td>
<td>-0.000002</td>
<td>0.000002</td>
<td>-0.918302</td>
<td>0.358759</td>
</tr>
</tbody>
</table>
ARMA(p,q) process

\[ y_t = a_0 + \sum_{k=1}^{p} a_k y_{t-k} + \sum_{l=1}^{q} b_k u_{t-k} + u_t \]

Combines autorregressive models with moving average models

Simple linear time-series model
Fitting to an ARMA(1,1)

timeseries: y
Method: Nonlinear Least Squares (Levenberg-Marquardt)
date: 03-03-10 time: 18:52
included observations: 755
p = 1 - q = 1 - constant - manual selection

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>4.62735411</td>
<td>0</td>
<td>148.9024</td>
<td>0</td>
</tr>
<tr>
<td>AR(1)</td>
<td>0.986154258</td>
<td>0</td>
<td>159.9401</td>
<td>0</td>
</tr>
<tr>
<td>MA(1)</td>
<td>-0.110605985</td>
<td>0</td>
<td>-2.998961</td>
<td>0.002798377</td>
</tr>
</tbody>
</table>

R-squared: 0.965239
Mean dependent var: 4.628068
Adjusted R-squared: 0.965147
S.D. dependent var: 0.071188
Akaike info criterion: -5.791955
Schwarz criterion: -5.773571
Log likelihood: 2189.462984
Durbin-Watson stat: 2.007356

Inverted AR-roots: 0.99
Inverted MA-roots: 0.11
Fitting y to an AR(1) process

timeseries: y

Method: Nonlinear Least Squares (Levenberg-Marquardt)
date: 03-03-10 time: 18:49
Included observations: 755
p = 1 - q = 0 - constant - manual selection

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>4.627528</td>
<td>0.03</td>
<td>168.4630632</td>
</tr>
<tr>
<td>AR(1)</td>
<td>0.98229241</td>
<td>0.01</td>
<td>143.6624447</td>
</tr>
</tbody>
</table>

R-squared          0.964800       Mean dependent var          4.628068
Adjusted R-squared 0.964753       S.D. dependent var          0.071188
S.E. of regression 0.013365       Akaike info criterion       -5.782052
Sum squared resid   0.134501       Schwarz criterion          -5.769796
Log likelihood     2184.724802     Durbin-Watson stat        2.225006
AR(1) coefficient for y estimated over a 60-day period

Red = upper bound for MR in 10 days, Green = upper bd for MR in 5 days
Dickey-Fuller over Sep 2008/March 2009

Augmented Dickey-Fuller test statistic  
-2.593218 0.284178

Test critical values:  
1% level -4.027516  
5% level -3.443485  
10% level -3.146482

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>tseries(-1)</td>
<td>-0.113728</td>
<td>0.043856</td>
<td>-2.593218</td>
<td>0.010671</td>
</tr>
<tr>
<td>D(tseries(-1))</td>
<td>-0.111532</td>
<td>0.090621</td>
<td>-1.230747</td>
<td>0.220785</td>
</tr>
<tr>
<td>D(tseries(-2))</td>
<td>0.162647</td>
<td>0.087448</td>
<td>1.859935</td>
<td>0.065303</td>
</tr>
<tr>
<td>D(tseries(-3))</td>
<td>0.040018</td>
<td>0.088750</td>
<td>0.450911</td>
<td>0.652854</td>
</tr>
<tr>
<td>D(tseries(-4))</td>
<td>0.267631</td>
<td>0.085738</td>
<td>-3.121501</td>
<td>0.002247</td>
</tr>
<tr>
<td>D(tseries(-5))</td>
<td>0.076574</td>
<td>0.086639</td>
<td>0.883828</td>
<td>0.378528</td>
</tr>
<tr>
<td>D(tseries(-6))</td>
<td>-0.139433</td>
<td>0.085911</td>
<td>-1.623007</td>
<td>0.107169</td>
</tr>
<tr>
<td>D(tseries(-7))</td>
<td>-0.242743</td>
<td>0.082689</td>
<td>-2.935598</td>
<td>0.003980</td>
</tr>
<tr>
<td>D(tseries(-8))</td>
<td>0.090026</td>
<td>0.085786</td>
<td>1.049428</td>
<td>0.296056</td>
</tr>
<tr>
<td>D(tseries(-9))</td>
<td>0.189077</td>
<td>0.084225</td>
<td>2.244910</td>
<td>0.026575</td>
</tr>
<tr>
<td>D(tseries(-10))</td>
<td>-0.106441</td>
<td>0.084083</td>
<td>-1.265896</td>
<td>0.207962</td>
</tr>
<tr>
<td>C</td>
<td>0.511426</td>
<td>0.199365</td>
<td>2.565270</td>
<td>0.011521</td>
</tr>
<tr>
<td>@trend</td>
<td>0.000090</td>
<td>0.000044</td>
<td>2.058864</td>
<td>0.041636</td>
</tr>
</tbody>
</table>
AR-1 coefficient for the period Sep 2008/march 2009

timeseries: ebay/xlk
Method: Nonlinear Least Squares (Levenberg-Marquardt)
date: 03-03-10 time: 18:40
Included observations: 145
p = 1 - q = 0 - constant - manual selection

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>4.55489463</td>
<td>0.013841</td>
<td>329.0952386</td>
<td>0</td>
</tr>
<tr>
<td>AR(1)</td>
<td>0.88029659</td>
<td>0.035204</td>
<td>25.00582423</td>
<td>-2.2E-16</td>
</tr>
</tbody>
</table>

R-squared 0.813873  Mean dependent var 4.558974
Adjusted R-squared 0.812571  S.D. dependent var 0.045858
S.E. of regression 0.019853  Akaike info criterion -4.952615
Sum squared resid 0.056364  Schwarz criterion -4.911557
Log likelihood 361.064616  Durbin-Watson stat 2.312544
Conclusions

ARCH, GARCH: models for volatility of financial series.

Volatility analysis via ARCH and GARCH lead to exponential moving averages of squared returns.

The advantage of GARCH over a fixed window is that GARCH is endogenous. However, fixed estimation windows for volatilities and correlations or exogenous EWMAs also make sense from a risk-management perspective.

Cointegration of stock prices via pairs is not easy to establish econometrically.

Unit root test: tests for stationarity

ARMA, AR: models for mean-reversion